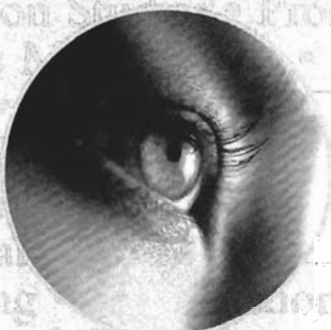


RESEARCH REPORT

DE LEGIOLEASE FILES  
EPISODE 4: HET BIED/LAAT-MYSTERIE

PIET SERCU

OR 0409



## *De Legiolease files*

### **Episode 4: het bied/laat-mysterie**

Piet Sercu\*

Eerste versie: Januari 2004; deze versie: Maart 2004



### **Abstract**

Het *Radar*-rapport over de aan- en verkoopkoersen gehanteerd door Legiolease is erg beïnvloed door het preferente aandeel Unilever, dat qua gedrag erg afwijkt van de rest en in realiteit bijlange niet het gewicht verdient die het in de studie krijgt. Ook brengen grote orders een prijsdruk mee van ongeveer een half procent, en het lijkt erg aannemelijk dat vooral beleggers met grote posities (en dus grote verliezen) op *Radar*'s oproep ingegaan zijn. Die twee overwegingen kunnen de uiteenlopende konklusies met het tweede rapport over de Legiolease-afrekeningen verklaren. Het lijkt me onwaarschijnlijk dat Dexia de gegevens voor het tweede rapport zou gemanipuleerd hebben. De konklusies van beide rapporten blijven onveranderd als je corrigeert voor afhankelijkheid tussen de afrekenkoersen over de aandelen heen.

# *De Legiolease files*

## Episode 4: het bied/laat-mysterie

### Wat voorafging:

**Episode 1.** In de wilde beursjaren '90 bood de Nederlandse bank Labouchère gehefboomde portefeuilles aan, zoals de "winstverdrievoudiger". De eenvoudigste manier om te hefboomen bestaat erin om naast de (bv.) EUR 100 van de belegger ook nog eens EUR 200 te ontfemen, zodat EUR 300 kan ingezet worden op de beurs ("leasing" van aandelen). Stijgt de beurs dan met 33,33 procent, dan groeit de aandelenpositie van 300 naar 400, zodat er na aflossing van de lening 200 overblijft voor de belegger zelf—een rendement van 100 procent op de eigen inzet van 100. Maar zo'n eenvoudige hefboom is onvermijdelijk ook een verliesverdubbelaar. Valt de beurs namelijk met 33,33 procent, dan zakt de aandelenportefeuille van 300 naar 200, zodat na aflossing van de lening niets meer overblijft voor de klant: het rendement wordt -100 procent.<sup>1</sup> Om rendementen beneden -100 procent te vermijden en de neerwaartse hefboom in het algemeen af te zwakken kan je bv put-opties toevoegen, maar het volledig wegwerken van de verliesverveelvoudiging leek prohibitief duur en overigens overbodig in een opgaande beurs. Het vervolg laat zich raden: met het ineenklappen van de koersen, vanaf 2000, eindigde het avontuur voor heel wat beleggers in tranen.

**Episode 2.** Rechtszaken worden aangespannen over de vraag of Labouchère, ondertussen overgenomen door Dexia, wel genoeg op de neerwaartse risico's heeft gewezen bij de klantenwerving. Het TV-programme *Radar* van de TROS wijdt een uitzending aan het Legiolease-debacle, en roept beleggers op om hun ervaringen voor te leggen. *Radar* ontvangt 150 afrekeningen, en speelt die door aan een Niet Nader Genoemde Onderzoeker (NNGO) aan de Erasmus-universiteit. De gegevens leiden tot een eindwerk van een student, waarin de afrekenprijzen vergeleken worden met de hoogste, laagste, en middenkoers van de dag. De koersen waartegen Legiolease-klanten kochten bleken gemiddeld een heel eind boven de middelkoers te liggen, en de verkoopkoersen een heel eind eronder. Het verschil is statistisch ondubbelzinnig groter dan een bied/laat-marge van 10 cent, en kan volgens de NNGO niet redelijkweze door een normale bied/laat-marge verklaard worden. De verdenking ontstaat dat Labouchère dus ook gesjoemeld zou kunnen hebben met de afrekenkoersen. *Radar* rapporteert over het verslag en plaatst het op hun webstek.

**Episode 3.** Dexia, ongelukkig met de negatieve publiciteit en mogelijks ook beducht om eventuele bijkomende rechtszaken, vraagt de Erasmus-Universiteit een grootscheepser onderzoek uit te voeren. De universiteit geeft die taak in handen van Wel Nader Genoemde Onderzoekers (WNGO), de econometristen Franses en Verbeek. Die ontvangen van Dexia meer dan 2100 afrekeningen, met de aankopen uit een *bull*-periode en verkopen uit een *bear*-episode. Ze onderzoeken pro Deo de aan- en verkoopkoersen. Die liggen volgens de WNGO helemaal niet abnormaal ver van elkaar. In hun reacties op het tegenrapport stellen *Radar* en *Vrij Nederland* vragen over de neutraliteit en onafhankelijkheid van de WNGO, omdat heel wat Erasmus-onderzoekers samenwerkingsverbanden hebben met o.m. Dexia. Of heeft Dexia de gegevens misschien opgeschoond? De episode eindigt met de vraag wie gelijk heeft, de NNGO dan wel de WNGO—of wellicht beiden?

**Indien u het begin van Episode 4 gemist hebt:** Ondergetekende hoort over de heibel in *Vrij Nederland* en leest beide rapporten na. Enige wachttijd in luchthavens wordt besteed aan het her-intikken van de *Radar*-gegevens, die afgeprint staan op het webstek-rapport. Het eindresultaat is onderhavige tekst.

---

<sup>1</sup>De voorbeelden verwaarlozen de leenrente. Het is een koud kunstje om na te gaan dat, rekening houdend met interest, de rendementen bovenop de leenrente verdrievoudigd worden ipv de bruto-rendementen zelf.

## Inleiding

De variabele waar het in deze discussie om gaat is de transactieprijs bij aan- of verkopen door Legiolease, relatief tegenover de hoogste en laagste koers van de dag. Als op een bepaalde dag de koersen bewogen tussen 100 en 104, en Legio bv. verkocht aan 102 of kocht aan 103, dan worden de transactiekoersen omgerekend tot scores van 0.5 (50%) en 0.75 (75%) op de hoogste/laagste-lat van die dag. Een nul-score komt dus overeen met een transactie aan de laagste dagkoers, score 1 (100 procent) aan de hoogste. Ik verwijs naar zo'n cijfer hierna als de relatieve koers. *Radar's* opinie is dat zowel de relatieve aankoopkoersen als de de relatieve verkoopkoersen abnormaal ver van de middenkoers (0,50) afliggen. Het tweede rapport vindt van niet.

Mijn bedenkingen zijn zuiver gebaseerd op de eerdere teksten geschreven, respectievelijk, in opdracht van *Radar* ("rapport 1") en op verzoek van Dexia ("rapport 2"). De gegevens gebruikt in rapport 2 zijn spijtig genoeg niet publiek beschikbaar. Die van Rapport 1 zijn dat wel: ze staan vermeld in bijlage van het rapport, dat van de *Radar*-webstek afgehaald kan worden. Ik heb zoals vermeld die cijfers terug ingetikt om bijkomende testen te kunnen uitvoeren, maar als bijproduct kwam er wel uit dat de berekeningen in rapport 1 volledig kloppen. De onderliggende gegevens van rapporten 1 en 2 heb ik niet direct kunnen toetsen aan de realiteit: dat kan alleen gebeuren via een grootschalige professionele audit, ter plaatse, van de informatiesystemen van Legio en Euronext.

Ik vat eerst kort beide rapporten samen; daarna bespreek ik potentiële gegevensproblemen (sectie 2) en methodologische aspecten (sectie 3).

## 1 Twee konflikerende diagnoses.

Tabel 1 geeft enkele kerncijfers uit beide rapporten. De gemiddelde relatieve koersen uit het eerste rapport liggen veel verder van de middenkoers (50%) af dan die uit het tweede rapport. De verschillen tussen de gemiddelden van rapporten 1 en 2 lijken significant (dwz kunnen niet redelijkerwijs nog aan het toeval toegeschreven worden) op basis van een konventionele t-test. Zo'n standaard t-test bouwt op de veronderstelling van onafhankelijkheid, maar de verschillen blijven significant na een benaderende correctie voor onderlinge afhankelijkheden.<sup>2</sup>

Een tweede (en totnutoe niet becommentarieerd) verschil betreft de varianties voor aan- versus verkooptransacties binnen het *Radar*-staal. In vergelijking met het tweede rapport<sup>3</sup> blijkt de variantie in *Radar's* verkoop-staal eerder groot, en die in het aankoop-staal bijzonder

---

<sup>2</sup>De correctie wordt beschreven en verantwoord in paragraaf 3.2.

<sup>3</sup>Tabellen 1 en 2 in dat rapport geven blijkbaar de standaarddeviaties van de gemiddelden, die ik dus kwadrateer en vermenigvuldig met de staalgrootte. De afgeprinte cijfers in die tabellen zijn erg onprecies—er is slechts 1 betekend cijfer opgegeven—maar de t-ratio en het gemiddelde impliceren een preciezer cijfer, dat hier gebruikt is.

Tabel 1: Enkele kenmerken van de twee stalen

	aankopen		verkopen	
	rapport 1	rapport 2	rapport 1	rapport 2
gemiddelde relatieve koers	0.65	0.54	0.32	0.51
scheefheid	-0.080	?	-0.579	?
variantie der relatieve koers	0.023	0.045	0.065	0.056
aantal observaties	64	683	61	1440
t-test verschil gemiddelden: klassiek		3.95		-5.67
idem, gecorrigeerd voor correlatie		2.77		-3.97

**Legenda.** Gemiddelden, varianties, en observatie-aantallen zijn afkomstig van de beide rapporten. De eerste t-test onderaan de tabel is de klassieke test op een verschil van twee gemiddelden met verschillende varianties; de tweede test infleert de varianties om benaderend te corrigeren voor afhankelijkheid tussen de relatieve afrekenkoersen van verschillende aandelen op eenzelfde dag; de details komen aan bod in paragraaf 3.2.

klein. Ik bereken voor staal 1 meteen ook de scheefheidscoëfficiënt, en vind alleen scheefheid in het verkoopstaal, met name naar links toe. Aan verkoopzijde zitten dus enkele uitzonderlijk kleine getallen.

In mijn discussie bekijk ik eerst verschillen in de gegevens: hoe zijn ze samengebracht, hoe waarschijnlijk is het dat er gesjoemeld is, en is alles redelijk representatief? Daarna bekijk ik de methodologische verschillen.

## 2 Een data-probleem?

Het eerste rapport werd uitgevoerd op vraag van TROS en werkt met 125 transacties, bekomen uit minstens 150 transacties vermeld in 48 contracten doorgegeven door TROS-kijkers na een *Radar*-oproep.<sup>4</sup> Het tweede rapport werd uitgevoerd op vraag van Dexia en werkt met 2123 transacties doorgegeven door Dexia. Een ander verschil is dat het tweede onderzoek de gegevens elimineert die slaan op het preferente aandeel Unilever. Dit effect werd namelijk tijdens de looptijd van de beleggingskontrakten als bonusaandeel uitgekeerd en heeft dus (i) een heel dunne markt (vandaar abnormaal slechte afrekenkoersen) en (ii) erg weinig gewicht in de totale af te wikkelen positie. De belangrijke verschillen zijn dus de bron der gegevens, de aantallen observaties, en het preferente aandeel Unilever.

Van *Radar* kreeg de tweede ploeg ook een aantal afrekeningen die niet in het eerste rapport gebruikt waren; dit maakt een steekproefsgewijze controle van de Dexia-gegevens mogelijk.

Mbt de uiteenlopende gemiddelden werden in de discussies twee mogelijke verklaringen geuit. Uit de *Radar*-website ("Vragen van TROS *Radar* aan EUR [12-06-2002]") blijkt de vrees dat Dexia de gegevens voor het tweede rapport misschien wel opgeleukt heeft. Rapport

---

<sup>4</sup>De getallen 48 contracten en 150 transacties komen voor in paragraaf 1 alinea 1; het cijfer 29 kontrakten in paragraaf 2, alinea volgend op het kadertje met de testvragen; en het cijfer 125 transacties uit de derde bijlage.

2 daarentegen denkt dat de TROS-oproep tot een niet-representatief staal geleid heeft. Ik bespreek beide en formuleer dan mijn eigen hypothese.

## 2.1 Kan Dexia de gevens van rapport 2 opgeschoond hebben?

Het lijkt me hoogst onwaarschijnlijk dat de basisgegevens geleverd door Dexia zouden kunnen vervalst of gefilterd geweest zijn zonder bij toetsing aan de *Radar*-gegevens door de mand te vallen. Ik redeneer bij mijn argumentatie *ex absurdo* vanuit de werkhypothese dat Dexia de gegevens vervalste. De bank leverde voor het tweede rapport alles bijeen 2123 contractprijzen. Zes van die records kwamen ook in het eerste onderzoek voor, dus kon Dexia nog maximaal  $2123 - 6 = 2117$  records bedokteren. De tweede ploeg had van *Radar* echter nog minstens 19 ( $= 48 - 29$ ) extra contracten gekregen, waarvan de datums Dexia niet bekend waren. Het aantal bijkomende transactie-records opgenomen in die 19 contracten is mij niet duidelijk, maar ik werk met drie kandidaat-cijfers: 98, 50, en 25. Het cijfer 98 is gebaseerd op de volgende redenering. Als de 19 ongebruikte contracten relatief gezien evenveel transacties bevatten als het eerste stel van 29, dan was er dus een controlestaal van  $150/29 \cdot 19 = 98$  transacties beschikbaar die Dexia niet kende. Andere cijfers in rapport 1 suggereren dan weer dat het staal veel kleiner kan geweest zijn. Het *Radar*-rapport vermeldt met name alles bijeen 150 transacties waarvan er 125 onderzocht werden, dus minstens 25 overblijvende transacties implicerend. Voor de goede orde bekeek ik ook het tussenliggend cijfer, 50.

Stap 1 in het argument is dat Dexia ruim boven 400 records vervalst zou moeten hebben om de verschuivingen van de gemiddelden te bekomen die we observeren. Vertrekpunt is het feit dat de variabelen logischerwijs en ook feitelijk in het interval  $[0, 1]$  liggen. (De histogrammen in rapport 2 bevestigen indirect dat dit nagekeken is.) Stel dat het echte gemiddelde voor de aankoopkant ongeveer 0.65 is, zoals het eerste rapport suggereert. Stel verder dat Dexia 200 records vervalste, gespreid over aan- en verkopen in proportie met het aantal aan- en verkopen, dus 66 vervalste aankopen en 134 vervalste verkopen. Stel verder dat de 66 aankopen, met oorspronkelijk gemiddelde 0.65, vervangen worden door nullen. Zelfs bij deze onrealistisch flagrante manipulatie daalt het gemiddelde amper van 0.65 naar  $(683-66)/683 \cdot 0.65 = 0.59$ . De maximale impact van 200 vervalsingen op het aankoopcijfer is dus amper  $0.65 - 0.59 = 0.06$ , waar wij bijna het dubbele zouden moeten zien. (Aan de verkoopkant is de vereiste impact nog groter.) Om een impact van 0.10 te hebben zouden 350 records op 0 of 1 moeten gezet zijn, of 700 op gemiddeld 0.33 of 0.67 enz.

Stap 2 in het argument is dat zelfs 200 vervalsingen vrijwel zeker ontdekt zouden worden zelfs bij amper 25 controles. Stel dat Dexia 200 op de 2117 records bedokterde. Als er 98 kontrolerecords zijn, is de kans dat een record veilig is (want niet bekend bij de WNGO) gelijk aan  $1 - 98/2117 = 0.95$ , maar de kans dat men 200 keer ongemerkt kan vervalsen daalt tot benaderend  $(1 - 98/2117)^{200} = .000076$ . Bij 50 of 25 kontrolerecords ipv 98 blijven de kansen nog steeds klein:  $(1 - 49/2117)^{200} = 0.0065$  en  $(1 - 25/2117)^{200} = 0.092$ . In Tabel 2 geef ik analoog de kans om *niet* tegen de lamp te lopen te worden als Dexia  $N = 100, 150, \dots, 400$  records (op 2117) manipuleert. Zelfs als er maar 25 kontrolerecords zijn blijft de kans om



**Tabel 2: maximale impact van  $N$  vervalsingen op het gemiddelde, en kans dat bij  $M$  controles géén van  $N$  veronderstelde manipulaties ontdekt wordt**

	N =							
	50	100	150	200	250	300	350	400
maximale impact	0.015	0.030	0.046	0.061	0.076	0.091	0.107	0.122
kans bij M= 25	0.548	0.296	0.158	0.082	0.051	0.028	0.016	0.009
kans bij M= 50	0.298	0.086	0.024	0.007	0.003	0.001	0.000	0.000
kans bij M= 98	0.091	0.008	0.001	0.000	0.000	0.000	0.000	0.000

**Legenda.** Er worden diverse theoretisch denkbare aantallen gemanipuleerde records bekeken ( $N$ ). Voor elke  $N$  wordt eerst nagerekende wat de impact zou zijn op de gemiddelde koopkoers indien de relatieve koers keihard op 0 (voor aankopen) of 1 (voor verkopen) zou gezet zijn. Ik noem dit de maximale impact van  $N$  omdat in de praktijk niemand zo extreem zou durven vervalsen. Vervolgens toon ik, voor elke  $N$ , de kans dat geen enkele afwijking aan het licht zou komen indien  $M(=\{25, 50, 98\})$  records getoetst worden aan de van *Radar* afkomstige data. Ik gebruik de hypergeometrische (zonder teruglegging) tot en met  $N=200$ , waarna Excel er het bijltje bij neergooit en ik overstap naar de binomiale benadering.

ontdekt te blijven prohibitief hoog, zeker als je de fraude opvoert tot 350, het minimumniveau. Je begint niet aan een fraude die met minstens 98.4 procent kans verkeerd afloopt.

Dexia kan dus geen voldoende aantal records gemanipuleerd hebben als er 98 of 49 of zelfs maar 25 controles waren. Ik konkludeer dat Dexia, bij 25 of meer controles, nooit onontdekt genoeg records kan bedokterd hebben om het gemiddelde te doen verschuiven in die mate die we observeren.

Dexia kan theoretisch ook de gegevens gefilterd hebben, door bv. het meest ongunstige derde van de records weg te laten. Maar ook dat moet onmiddellijk aan het licht gekomen zijn bij toetsing aan de *Radar*-gegevens. Stel dat er oorspronkelijk 3117 records waren, en dat Dexia daarvan het top-derde (1000 observaties) weglaat zodat het gemiddelde daalt van 0.65 naar 0.50. De kans dat geen van die 1000 weggelaten records in het kleinste controlestaal voorkomt is dan benaderend  $(1 - 25/2117)^{1000} = 0.000006$ . Het gebruik van het controlestaal moet dus volstaan om zowel vervalsingen als weglatingen door Dexia te ontdekken, omdat elke vervalsing of weglating op zeer massale schaal zou moeten gebeurd zijn om de verschuivingen te kunnen opleveren die we gezien hebben.

## 2.2 Zelfselectie-vertekening in het *Radar*-staal op basis van de afrekenprijs?

In de bijlage van het tweede rapport wordt omgekeerd de mogelijkheid van zelfselectie-vertekening in het *Radar*-staal geopperd. Beleggers met grotere verliezen zijn meer geneigd op de TROS-oproep in te gaan, en een ongunstige afrekenkoers in vergelijking tot het daggemiddelde draagt wel degelijk bij tot een groter verlies. Wie zich aanbiedt heeft dus waarschijnlijk onrepresentatief ongunstige relatieve koersen gekregen.

Dit mechanisme is logisch onweerlegbaar maar waarschijnlijk in de praktijk verwaarloosbaar in vergelijking met de totale koersdaling. De kleinste koersdaling, gemeten tussen de eerste aankoop en de laatste verkoop vermeld in het *Radar*-rapport, is -56 procent (Unilever), de grootste -89 procent (KPN). Een extra ongunstige afrekening ten belope van 0.10 van de

hoog/laag-marge (die 3.43 procent van de koers bedraagt) betekent één derde procent extra verlies, hetgeen marginaal is binnen het totale verlies van 56 tot 89 procent.

### 2.3 Zelfselectie-vertekening in het *Radar*-staal op basis van de ordergrootte?

Er kan in het *Radar*-staal een andere vorm van zelfselectie-vertekening binnengeslopen zijn, namelijk via de tweede determinant van het totale verlies, de grootte van de positie. Grote posities afwikkelen brengt prijsdruk mee, en dus ongunstige afrekenkoersen.

Mijn uitwerking van dit idee is gebaseerd op de microstructuur van Euronext. Beleggers delen aan Euronext vóór de opening hun limietkoersen voor aan- of verkoop mee, dwz hun vraag- en aanbodfuncties. Euronext aggregeert dit tot marktvraag en -aanbod, en kruist bij de opening vraag en aanbod (de call-markt). Daarna blijven de niet-uitgevoerde aan- en verkooporders zichtbaar op de Euronext-schermen ("het limietorderboek"). Iedereen die na de opening-call wil (ver)kopen kan dus openstaande orders opnemen, of zijn/haar limietorders aan het boek toevoegen. De bid-ask-spread of prijsvork is hier dus het verschil tussen de prijzen in het beste nog uitstaande verkooporder en het beste nog uitstaande aankooporder. Als een beursagent een groot order binnenkrijgt mag hij/zij die buiten het boek om uitvoeren. Rapport 1 vermeldt de regels hieromtrent, en bevindt dat de transacties van Legiolease te klein zijn om in die categorie te kunnen vallen,

Rapport 2 spreekt niet over prijsdruk. Rapport 1 onderkent het probleem van prijsdruk maar stelt zich tevreden met de vaststelling dat de kwantiteiten niet groot genoeg zijn om buiten beurs en buiten de prijsvork te mogen gaan. Daarmee is de kous echter niet af. Prijsdruk, binnen Euronext, ontstaat als het order groter is dan de kwantiteit die uitstaat aan de beste prijs (de vorkprijs): dan moet de koper twee of meer niveaus het boek inklimmen (of moet de verkoper twee of meer niveaus in het boek afdalen) om genoeg kwantiteit bijeen te harken. In dat geval is de gemiddelde transactiekoers slechter dan de vork-koers. Hoe groter de af te wikkelen hoeveelheid, des te verder men het boek moet ingaan en hoe ongunstiger dus de gemiddelde prijs. Het is kortom, in tegenstelling tot in een angelsaksische "market-maker" organisatie, niet duidelijk wat de normale bied/laat-marge is: (i) de breedte van de vork varieert min of meer lukraak doorheen de tijd, en (ii) voor heel wat transacties moet men "het boek ingaan" en dus aan slechtere prijzen werken dan de vork-koersen.

Het is wél gemakkelijk na te gaan of het order slechts één niveau in het limietorderboek moest aanboren dan wel twee of meer. Prijzen moeten namelijk een veelvoud zijn van een basisnorm, de *tick*-grootte. De *tick* voor prijzen tussen 100 en 500, bvb, is 0.10, wat betekent dat die prijzen eindigen op een veelvoud van 0.10. Wanneer de Legioprijs geen veelvoud is van de *tick*-grootte, is die prijs noodzakelijk een gemiddelde van meerdere prijzen, wat betekent dat Legio twee of meer niveaus heeft moeten aanspreken om het order te kunnen uitgevoerd krijgen. Ik bekijk dus apart de relatieve koersen voor gevallen waar de prijs sloeg op één niveau versus op twee of meer niveaus. Dit werkt echter niet bij koersen beneden 50: beneden 50 Euro is de *tick*-grootte 0.01, en omdat ook gemiddelde prijzen afgerond zijn op cents zien we dus

**Tabel 3: regressie-analyse van de impact van grote versus kleine orders**

$$\frac{P_{i,t}^{Legio} - P_{i,t}^{low}}{P_{i,t}^{high} - P_{i,t}^{low}} = \alpha + \beta I_{it}^{buy} + \Delta_{\alpha} I_{i,t}^{sell+avg} + \Delta_{\beta} I_{i,t}^{buy+avg} + \epsilon_{i,t} \quad (2.1)$$

	relatieve koers voor kleine orders		extra afwijking voor grote orders	
	aankoop	verkoop	aankoop	verkoop
impliciete gemiddelde relatieve koers	0.594	0.474	0.656	0.327
coëfficiënt zoals geschat in regressie	+0.120	0.474	+0.062	-0.147
stdev	0.080	0.059	0.061	0.075
t-test tov $H_0$ : gem. verkoopprijs=0.50	—	-0.448	—	—
t-test tov $H_0$ : coefficient=0	1.495	—	1.029	-1.941
aantal observaties	43	11	21	10

$$R^2 = 0.35, F(4,81)=14.07, p=0.000$$

Het staal bevat de 85 prijzen boven 50 Euro, waar het verschil tussen gemiddelde en niet-gemiddelde prijzen zichtbaar is. De regressand is de aan- of verkoopkoers gerapporteerd door Legio, gesitueerd relatief tegenover de hoogste en laagste dagkoers.  $I_{it}^{buy} = 1$  als het om een aankoop gaat, anders 0;  $I_{i,t}^{sell+avg} = 1$  als het om een verkoop aan een gemiddelde koers gaat (groot verkooporder) en  $I_{i,t}^{buy+avg} = 1$  als het om een aankoop aan een gemiddelde koers gaat (groot aankooporder). De coëfficiënt  $\alpha$  meet dus de typische prijs voor een klein verkooporder;  $\beta$  meet het verschil tussen de gemiddelde transactiekoersen voor een klein aan- en verkooporder, zodat  $\alpha + \beta$  de typische relatieve koers oplevert voor een kleine aankoop;  $\Delta_{\alpha}$  en  $\Delta_{\beta}$  meten hoeveel de verkoop- of aankoopkoersen voor grote orders afwijken van die voor kleine orders.

geen verschil meer tussen gemiddelde en gewone koersen. Dit laat ons 85 observaties, waarvan 54 "groot". Tabel 3 toont dat de grote transacties iets vaker voorkomen bij aankopen (43 van de 64) dan bij verkopen aan prijzen boven 50 (11 op de 21). Verder blijkt dat voor kleine transacties de gemiddelde relatieve aan- en verkoop algebraïsch en statistisch dichter bij 0.50 te liggen dan de algemene cijfers: bij aankoop 0.59 ipv 0.65, bij verkoop 0.47 ipv 0.32. Een flink stuk van de actie komt dus van de grote transacties: aan grote aankoop kost naar schatting 0.062 extra (zij het statistisch niet significant), een grote verkoop brengt naar schatting 0.147 minder op (statistisch significant).<sup>5</sup> Dit is consistent met de prijsdrukhypothese bij verkoop—die gebeurde in een neergaande, dunne markt—en mogelijks ook bij aankoop, al is dat effect uiteraard niet zo duidelijk merkbaar in een stijgende, liquide markt. De prijsdruk verklaart dus, via zelfselectie, mogelijks het verschil tussen de twee studies.

## 2.4 Het preferent bonusaandeel Unilever

Een ander verschil tussen beide stalen is dat het tweede onderzoek de gegevens elimineert die slaan op het preferente aandeel Unilever. Dit effect werd namelijk tijdens de looptijd van de beleggingskontrakten als bonusaandeel uitgekeerd en heeft dus een heel dunne markt. Het is

<sup>5</sup>De marge hoog-laag is volgens rapport 2 gemiddeld 3.43%. Een minderopbrengst van 0.147 van de hoog/laag-marge komt dus overeen met 0.5% van de prijs. De geschatte prijsdruk bij de aankopen was minder dan half zo groot.

goedkoop, ongeveer 6 euro, en heeft dus erg hoge transactiekosten en een grote relatieve *tick*-grootte. Het P-aandeel is ergens 5% waard van het gewoon aandeel waarvan het afgesplitst is, en moet dus minder dan 1% van de transacties (in waarde) vertegenwoordigen. Toch staat het voor 10 (op 61) verkooptransacties in het *Radar*-staal.

Het effect wordt blijkbaar gemiddeld verkocht aan een erg ongunstige koers, vermoedelijk dus het gevolg van prijsdruk in een erg dunne markt met ongewoon grote relatieve *tick*-grootte. Die afrekenkoersen blijken ook een variantie te hebben die twee keer hoger ligt dan bij de overige verkoopkoersen, om dezelfde redenen. Als ik de gegevens m.b.t. dit onrepresentatieve aandeel uit het *Radar*-staal weglaat, dan verdwijnen twee anomalieën in rapport 1, met name de asymmetrie der gemiddelden en de grote variantie en linksscheefheid aan de verkooptant:

	totaal staal	Unilever P-aandeel	rest van staal
gemiddelde (verkoop)	0.319	0.049	0.373
variantie (verkoop)	0.063	0.085	0.043

## 2.5 Besluit m.b.t. de gegevens

Samengevat: a priori lijkt het waarschijnlijk dat er in Staal 1 ongewoon veel grote posities zaten, met prijsdruk als gevolg. De kleinere afrekeningen lijken inderdaad erg op die van Rapport 2, terwijl grote posities ergens een half procent nadeliger uitvallen. Het preferent aandeel Unilever, anderzijds, heeft een uitzonderlijk slechte record en krijgt een overmatig gewicht in Staal 1. Die twee factoren verklaren voor een flink stuk de verschillen tussen beide rapporten.

Toch is niet alles opgelost. De kleine variantie van de aankoopkoersen in het *Radar*-staal blijft mij een raadsel, zoals ook de gunstige gemiddelde verkoopkoers in staal 2. Een derde vraagteken is de éne waarneming, in rapport 1, van een afrekenig (preferent aandeel Unilever) die buiten de hoogste/laagste-marge van de dag ligt.

Na de bespreking van verschillen in de gegevensverzameling bekijk ik de methodologische aspecten.

## 3 Testprocedures en konklusies in beide rapporten.

### 3.1 De gevolgde procedures

Ik begin met een korte beschrijving van de beide testprocedures. Rapport 1 test eerst of de gemiddelde relatieve transactiekoers voor aan- en verkoop wellicht aan elkaar gelijk kunnen zijn, en verwerpt dit. Vervolgens test dat rapport of de gemiddelde aan- en verkoopkoersen wellicht met niet meer dan 10 cent van elkaar kunnen verschillen (aannemende dat bied- en laatkoersen normaliter met 10 cent van elkaar afwijken), en verwerpt ook die hypothese.

Het moet me van het hart dat het testen van de hypothese van géén verschil neerkomt op het intrappen van een open deur. Een tweede betwistbaar punt of het redelijk is een normale

bied/laat-marge te poneren in cents, voor alle aandelen ongeacht hun prijsniveau. We bekijken dit verderop. Het niveau van die éne bied/laat-marge, 10 cent, lijkt me even onredelijk, zeker aan de aankoopzijde waar de prijs gemiddeld 115 euro is en de spread dus geen tiende van een percent zou belopen.<sup>6</sup> Maar goed: rapport 1 gaat vervolgens over tot het berekenen van een impliciete bied/laat-marge, dwz de hypothetische marge waarbij de verschillen wél aanvaardbaar zouden worden. De verschillen tussen relatieve aan- en verkoopkoersen worden volgens het *Radar*-rapport slechts statistisch aanvaardbaar bij een bied/laat-marge die "zeer onwaarschijnlijk" bevonden wordt. Het is spijtig genoeg niet duidelijk of die kritische bied/laat-marge moet geïnterpreteerd worden als een spread of een halfspread. In de tekst begin paragraaf 3 worden spreads en half-spreads namelijk manifest met elkaar verward—tekst en tabellen spreken mekaar tegen—zodat ik er het raden naar heb wat ze verderop bedoelen. Ik doe dat werk dus straks over, behalve dat ik een kritische spread in procenten zoek eerder dan een in centen.

Het tweede rapport test of de gemiddelde relatieve aankoopkoers afwijkt van de middenkoers (50%), en verwerpt dit. Als de normale bied/laat-marge 0.25 procent van de koers bedraagt wordt de afwijking van de aankoopkoers tov de middenkoers statistisch onduidelijk. De gemiddelde verkoopkoers daarentegen wijkt niet significant van de middenkoers af, en wordt zelfs statistisch uitgesproken klantvriendelijk als de normale bied/laat-marge 0.25 procent van de koers belooft.

Een eerste verschil tussen beide rapporten mbt de test-opzet heeft dus te maken met de basisvraag: moet je testen of gemiddelde aan- en verkoopkoersen elk apart afwijken van de middenkoers (50 procent), of moet je eerder het verschil testen tussen gemiddelde aan- en verkoopkoersen? Rapport 2 doet het eerste. Testen of gemiddelde relatieve aan- en verkoopkoersen elk afzonderlijk afwijken van 50% vereist een bijkomende veronderstelling die niet nodig is in de *Radar*-procedure, namelijk dat een lukraak gekozen transactie gemiddeld aan de middenkoers gebeurt. Die hypothese wordt wel ondersteund door de testresultaten voor de openings-, slot- en waardewogen dagkoers. Het voordeel is dat men eventueel anomalieën kan traceren naar de aan- of verkoopzijde. Al bij al lijkt deze procedure verdedigbaar, en zonder grote impact op het eindresultaat.

Een ernstiger methodologisch probleem is de aanname, in beide rapporten, dat de waarnemingen onderling onafhankelijk zijn. Beide onderkennen dit als een probleem. Omdat transacties gegroepeerd verlopen en koersen samen bewegen, ook binnen de dag, gaat een ongewoon hoge (of lage) koers voor bv. ING meestal samen met een eveneens hoge (of lage) koers voor Fortis. Die afhankelijkheid invalideert de standaard-testen. *Radar* doet daar verder weinig mee, en het tweede rapport biedt slechts een begin van oplossing. Ik bekijk eerst dit probleem, omdat de oplossing ook bij verdere berekeningen bruikbaar is. De belangrijkste daarna nog te bespreken verschillen zijn hier dus (*i*) de keuzen voor een procentuele dan wel absolute

---

<sup>6</sup>De gemiddelde prijs voor de verkopen daarentegen is amper 35 euro.

bied/laat-marge, en (ii) de vraag hoe groot een normale bied/laat-marge is.

### 3.2 Statistische (on)afhankelijkheid

Beide rapporten berekenen een *gepooled* gemiddelde over alle aandelen heen, en een klassieke t-test die statistische onafhankelijkheid veronderstelt. Auto- en kruiselingse correlatie kan geen groot probleem opleveren, maar gelijktijdige correlaties zijn erg belangrijk zoals blijkt uit de appendix van het tweede rapport: de coëfficiënten lopen op tot 0,50.

Het *Radar*-rapport exploreert een mogelijke impact van de afhankelijkheid op de teller van de t-test, het gemiddelde: als transacties stevast op (of rond) hetzelfde uur gebeuren *en* als er een intradagelijks seizoenspatroon is, kunnen we mogelijks zo'n seizoenspatroon aan het meten zijn eerder dan abnormale koersen. Deze hypothese wordt door hen niet bevestigd. Het tweede rapport focust op de betrouwbaarheid der t-test. Er wordt opgemerkt dat een aandeel-per-aandeel analyse betrouwbaarder is. Dit is maar ten dele waar: men krijgt weliswaar negen t-tests die elk op zich betrouwbaarder zijn, maar de negen cijfers blijven onderling afhankelijk zodat de inferentie uit het totale plaatje nog steeds niet meteen duidelijk is.

Alhoewel beide rapporten dus tekortschieten op dit gebied, is de impact duidelijk erger voor het eerste rapport. Zij vinden hoge t-statistics, en horen zich dus de vraag te stellen in welke mate de ratio's overschat zijn door het onderschatten van de noemer, de standaarddeviatie. De auteurs van het tweede rapport daarentegen eindigen ondanks de onderschatting der noemer reeds met kleine ratio's en kunnen zich dus tevreden stellen met de wetenschap dat een correctie de konklusie alleen maar kon versterkt hebben.

Het is vrij eenvoudig een ruwe grootte-orde van de maximale impact van de correlaties te berekenen. Ik maak drie veronderstellingen: (i) alle varianties zijn dezelfde over de aandelen heen, (ii) er zijn geen auto- of kruiselingse correlaties, en (iii) de contemporane correlaties zijn gelijk aan 0.5. De veronderstelling over de varianties lijkt *a priori* redelijk (de variabelen moeten ongeveer uniform zijn verdeeld over de zone 0-1) en werkt ook empirisch goed, zie appendix. Mijn notatie is:

$\bar{X}$	=	het gepoolde gemiddelde over aandelen $a = 1, \dots, A$ en dagen $d = 1, \dots, D$
$X_{a,d}$	=	de positie van de Legio-prijs $(a, d)$ binnen de hoogste/laagste-range voor $(a, d)$
$\sigma^2$	=	de (gemeenschappelijke) variantie
$N_d$	=	het aantal aandelen waarin Legio handelde op dag $d$
$N$	=	het totaal aantal observaties

Dan is de juiste variantie van het gepoolde gemiddelde gelijk aan<sup>7</sup>

$$\text{var}(\bar{X}) = \frac{\text{var}\left(\sum_{d=1}^D \sum_{a=1}^{N_d} X_{a,d}\right)}{N^2} \quad (3.2)$$

<sup>7</sup>De tweede lijn volgt uit de aanname van zero intertemporele correlaties. De derde lijn volgt uit het feit dat in de variantie van een som van  $N_d$  variabelen er  $N_d$  varianties steken en  $N_d^2 - N_d$  covarianties, de laatste ingeschat als de helft van de variantie omwille van de correlaties van om en bij 0.5.

**Tabel 4: Inschatting van het effect van correlatie op de variantie**

verdeling van het aantal aandelen		berekening incl. de covarianties		id. zonder covarianties
$N_d(i)$	$freq(i)$	$\frac{(N_d^2 + N_d)}{2}$	$id \times freq(i)$	$N_d \times freq(i)$
1	1	1	1	1
2	1	3	3	2
3	6	6	36	18
4	11	10	110	44
5	0	15	0	0
6	10	21	210	60
som	29	n.v.t.	360	125

**Legenda.**  $N_d$  stelt het mogelijk aantal transacties per dag voor, dat in rapport 1 uiteenloopt van één tot zes. Het aantal keer dat er inderdaad  $N_d$  aandelen verhandeld werden wordt aangegeven in de kolom  $freq$ . In de twee volgende kolommen wordt de variantie berekend van de som rekening houdend met de covarianties, en in de laatste kolom zonder daarmee rekening te houden (de standaardprocedure).

$$= \frac{\sum_{d=1}^D \text{var} \left( \sum_{a=1}^{N_d} X_{a,d} \right)}{N^2} \quad (3.3)$$

$$= \frac{\sum_{d=1}^D (N_d \sigma^2 + (N_d^2 - N_d) \sigma^2 / 2)}{N^2} \quad (3.4)$$

$$= \frac{\sum_{d=1}^D (N_d^2 + N_d) \sigma^2 / 2}{N^2} \quad (3.5)$$

Het *Radar*-rapport laat ons toe de variantie te begroten omdat we kunnen natellen hoe vaak er  $N_d(i) = 1, 2, \dots, 6$  aandelen per dag verhandeld werden.

De basiscijfers en de berekeningen vind je in Table 4. Het in rekening brengen van de covarianties kan de variantie van de *gepoolede* som dus doen stijgen van 125 keer  $\sigma^2$  naar ongeveer 360 keer. M.a.w., de echte variantie van het gemiddelde kan dus tot 360/125 keer groter zijn dan de in het eerste rapport gehanteerde variantie. Uit  $\sqrt{125/360} = 0.6$  volgt dat we dus de t-testen moeten doen krimpen met tot 40 percent tegenover de eerder berekende niveaus. Dit heb ik onderin Tabel 1 toegepast, en ik konkludeer dat de verschillen significant blijven.

### 3.3 De bied/laat-marge—centen of procenten?

Het eerste rapport werkt in eerste instantie met een spread van 0.10 Euro voor alle aandelen, ongeacht de koers. Dit klopt niet de organisatie van Euronext. Het best beschikbare equivalent, op Euronext, van de Anglosaxische bied/laat-marge is de prijsvork, het verschil tussen het laagste laatprijs onder alle openstaande limiet-verkooporders en de hoogste biedprijs onder alle openstaande limiet-kooporders. De vork is minstens een *tick* breed, en de grootte van die *tick* stijgt met de koers, zie Tabel 5. Maar de prijszones waarin de *tick*-grootte konstant blijft zijn erg breed. De situatie is dus, a priori, noch duidelijk een cents-verhaal (zoals rapport 1 veronderstelt) noch duidelijk een procenten-verhaal (rapport 2): je zou een tussensituatie kunnen hebben waarin er in elke *tick*-grootte-zone er een vaste gemiddelde vorkbreedte in cents

bestaat, dwz een stap-functie ipv een monotoon stijgende or een vlakke lijn.

**Tabel 5: tickgrootte, en frequentie der transacties, per prijsklasse**

prijsklasse	tick-grootte	als % van grens		frequentie van voorkomen, bij		
		onder	boven	aankoop	verkoop	pooled
Eur 0-50	Eur 0.01	(infinite)	0.02%	21%	67%	44%
Eur 50-100	Eur 0.05	0.1%	0.05%	13%	33%	23%
Eur 100-250*	Eur 0.10	0.1%	0.04%	67%	0%	34%
som				100%	100%	100%

\*: 250 is de hoogste feitelijke geobserveerde prijs; de klassegrens is 500.

A priori lijkt me de in rapport 1 gehanteerde spread wel erg laag, vooral aan het top-einde der prijzen. Aangezien de prijzen uiteenlopen van 5 tot (in dit staal) 250 Euro, betekent de veronderstelde vaste spread van 0.10 Euro een procentuele spread die zou variëren van zomaar eventjes 2 tot 0.02 procent. Voor de meeste aandelen liggen de prijzen in de range 50-150, en de begrote kosten dus op 0.10 tot 0.03 procent, wat me onrealistisch laag lijkt. Het rapport vermeldt ook geen bron voor hun aanname, maar doet wel een reeks experimenten om na te gaan bij welke spread de t-testen hun significantie verliezen. Dit gebeurt bij 0.60 Euro. Zoals vermeld is het niet duidelijk of dit een half-spread is dan wel een echte spread, maar zeker als het een half-spread is lijkt dit aan de hoge kant—aannemend dat het werken met een onveranderlijke spread (ongeacht de prijs) zin heeft.

Het tweede rapport poneert a priori zijn redelijk geachte cijfers, 0.25-0.50 procent van de koers, en citeert hiervoor gepubliceerd onderzoek. (Het levert ook een cijfer voor de gemiddelde *high/low spread*, 3.4 procent, dat in dit commentaar herhaaldelijk gebruikt wordt.) Een andere bron citeert, met wonderbaarlijke precisie, een (half-spread) kost van 0.245 procent voor Amsterdam.<sup>8</sup> Aktas en Van Oppens citeren voor enkele aandelen op Euronext-Parijs een spread van om en bij de 0.35 procent, gedefinieerd als het tijdsgewogen gemiddelde verschil tussen de beste bied- en laatkoers in het limietorderboek.<sup>9</sup> Acharya en Pedersen tonen Amerikaanse éénrichtings transactiekosten die uiteenlopen van 0.25 (meest liquid) procent naar 0.40 (mediaan) en tot 8.83 (laagste 4 procent qua liquiditeit).<sup>10</sup>

Intuïtief lijkt een procentuele kost redelijker dan een vaste kost in cents, en in de literatuur werkt men dan ook stevast met procenten eerder dan centen; maar een tussenscenario is denkbaar waarbij de trapsgewijze *tick*-grootte een trapsgewijze kostenstructuur induceert. Ik test dus eerst formeel of er in de relatieve transactiekoers een bid-ask-effect zit dat proportioneel is met de *tick*-grootte. Ik hypotheeser daarom dat de afwijking tegenover de middenprijs

<sup>8</sup>Domowitz, I., J. Glen & A. Madhavan, 2000, Liquidity, Volatility, and equity trading costs across countries and over time, Penn State U, Working Paper

<sup>9</sup>Aktas. N. en H. Van Oppens, Probability of Informed Trading? Some evidence around corporate events, working paper, Université Catholique de Louvain, 2004

<sup>10</sup>Acharya. V. en L. Pedersen, Asset pricing with liquidity risk, working paper, London Business School, 2003



Tabel 7: regressie-analyse van de impact van de *tick*-grootte
$$\frac{P_{i,t}^{Legio} - P_{i,t}^{low}}{P_{i,t}^{high} - P_{i,t}^{low}} = \alpha + \beta I_{it}^{buy} + \gamma \left[ \frac{tick-grootte(P_{i,t}^{Legio})}{P_{i,t}^{high} - P_{i,t}^{low}} \times I_{it}^{buy} \right] + \delta \left[ \frac{tick-grootte(P_{i,t}^{Legio})}{P_{i,t}^{high} - P_{i,t}^{low}} \times (1 - I_{it}^{buy}) \right] + \epsilon_{i,t}$$

	$\alpha$	$\beta$	$\gamma$	$\delta$
coëfficiënt	0.396	0.320	-0.957	-0.644
stdev	0.467	0.083	0.974	1.054
t-test tov $H_0: \text{coef}=0.50$	-2.218	—	—	—
t-test tov $H_0: \text{coef}=0$	—	3.868	-0.982	-0.611

$R^2 = 0.38, F=22.83, p=0.000$

**Legenda.** Het staal bevat de 115 prijzen boven 10 Euro (dus zonder het preferente bonusaandeel Unilever). De regressand is de aan- of verkoopkoers gerapporteerd door Legio, gesitueerd relatief tegenover de hoogste en laagste dagkoers.  $I_{it}^{buy} = 1$  als het om een aankoop gaat, anders 0; en *tick-grootte*( $P$ ) is de *tick*-grootte die hoort bij de prijs.

een lukraake component bevat—de beweging binnen de dag rond de middelprijs, hieronder aangeduid als  $\nu_{i,t}$ —plus een spreadcomponent die proportioneel is aan *tick-grootte*( $P$ ), de *tick*-grootte die hoort bij de prijs:

$$P_{i,t}^{Legio,buy} = P_{i,t}^{mid} + \gamma \cdot tick-grootte(P_{i,t}^{Legio}) + \nu_{i,t} \quad (3.6)$$

$$\Rightarrow \frac{P_{i,t}^{Legio,buy} - P_{i,t}^{low}}{P_{i,t}^{high} - P_{i,t}^{low}} = \frac{P_{i,t}^{mid} - P_{i,t}^{low}}{P_{i,t}^{high} - P_{i,t}^{low}} + \gamma \frac{tick-grootte(P_{i,t}^{Legio})}{P_{i,t}^{high} - P_{i,t}^{low}} + \frac{\nu_{i,t}}{P_{i,t}^{high} - P_{i,t}^{low}} \quad (3.7)$$

$$= 1/2 + \gamma \frac{tick-grootte(P_{i,t}^{Legio})}{P_{i,t}^{high} - P_{i,t}^{low}} + \epsilon_{i,t} \quad (3.8)$$

In die logica verwacht ik dus, in een cross-sectionele regressie van de relatieve transactiekoers op de geschaalde *tick*-grootte, een intercept van om en bij de 0.5, en een positieve gamma, te interpreteren als de helft van het aantal *ticks* tussen een lukrake aankoop en een lukrake verkoop binnen dezelfde dag. Een analoge vergelijking voor de verkoopprijzen is

$$\frac{P_{i,t}^{Legio,sell} - P_{i,t}^{low}}{P_{i,t}^{high} - P_{i,t}^{low}} = 1/2 + \delta \frac{tick-grootte(P_{i,t}^{Legio})}{P_{i,t}^{high} - P_{i,t}^{low}} + \epsilon_{i,t} \quad (3.9)$$

waar  $\delta$  negatief zou moeten zijn en gelijk aan  $-\gamma$ . Ik schat beide vergelijkingen samen, zoals getoond in Tabel 7. Ik loop dus een regressie op vier variabelen: een constante, een aankoop-dummy die op 1 staat als het om een aankoop gaat en op 0 als het een verkoop betreft, de relatieve *tick*-grootte die hoort bij elk van de aankoopkoersen, en de relatieve *tick*-grootte die hoort bij elk van de verkoopkoersen. De Unilever Preferent-observaties laat ik buiten beschouwing. Ik vind dat die *tick*-grootte-regressoren statistisch insignifkant zijn. Er is geen aantoonbaar positieve gamma en geen aantoonbaar negatieve delta: de tekens van de (insignifikante) schatting aan de aankoopzijde is zelfs omgekeerd aan de logische verwachtingen. De intercepten voor aan- en verkoop,  $\alpha$  en  $\alpha + \beta$ , liggen tot slot ver van 0.5 af. Ik besluit dat het bestaan van brede zones met daarbinnen een vaste *tick*-grootte geen aantoonbaar stap-

**Tabel 8: T-testen voor aan- en verkoopprijzen voor diverse normale procentuele spreads en met correctie voor correlatie**

	aankopen	verkopen
<i>Radars</i> gemiddelde	0.653	0.320
<i>Radars</i> variantie	0.025	0.064
<i>Radars</i> aantal observaties	64	61
t-test tov $E(X) = 0.5 \pm 0.30/3.4/2$	3.89	-2.93
t-test tov $E(X) = 0.5 \pm 0.40/3.4/2$	3.36	-2.61
t-test tov $E(X) = 0.5 \pm 0.50/3.4/2$	2.84	-2.30
t-test tov $E(X) = 0.5 \pm 0.60/3.4/2$	2.31	-1.98
t-test tov $E(X) = 0.5 \pm 0.70/3.4/2$	1.79	-1.66

patroon oplegt aan de relatieve transactiekoersen, hetgeen de simpele procentuele benadering van rapport 2 aanvaardbaar maakt.

### 3.4 De kritische procentuele kostenmarge in Staal 1

De resterende vraag is bij welke procentuele bied/laat-marge in het *Radar*-staal de verschillen tussen aan- en verkoopkoersen verklaarbaar worden. Gegeven een high/low marge van 3.43 % van de koers en een bied/laat-marge van  $p$  % van de koers, vertegenwoordigt de kost (de half-spread) een fractie  $p/3.43/2$  van de high/low-marge. Ik toon in Tabel 8 de resultaten van experimenten met  $p = \{0.3, 0.4, \dots\}$ . Om t-ratio's beneden 2 te verkrijgen moet ik  $p$  opvoeren tot 0.7 %, hetgeen ook de auteurs van rapport 2 hoog zouden gevonden hebben en zeker duurder uitvalt dan de 0.245% die ik bij Domowitz *et al.* aantref. Zuivere bid-ask-kosten in de zin van de prijsvork verklaren dus niet de geobserveerde kosten, en het *Radar*-staal wijkt wel degelijk af van het *Dexia*-staal ook na correctie voor verschillende berekeningswijzen en voor onderlinge afhankelijkheden.

## 4 Besluit

Mijn hoofdbevindingen zijn als volgt:

- Beide rapporten lijken me bona fide en ik zie geen gegronde redenen om aan de juistheid van de gegevens te twijfelen.

Stel in het bijzonder, en *ex absurdo*, dat het het eerste rapport het echte gemiddelde had opgeleverd en dat de gegevens in rapport in rapport 2 door *Dexia* opgeschoond zouden zijn: dan zou, om de verschuiving in de gemiddelden te krijgen die wij nu zien, *Dexia* minstens één record op drie hebben moeten vervalsen, hetgeen zelfs bij een relatief beperkt controle onvermijdelijk zou opgevallen zijn.

- Elk rapport had beter zijn best kunnen doen om de onderlinge onafhankelijkheden in rekening te brengen. Voor rapport 1 was de correctie meest nodig omdat dit de konklusies had kunnen doen omslaan; voor rapport 2 was dat niet zo.

- In het staal gebruikt in rapport 1 liggen de transactiekoersen significant verder van de middenkoers dan in het staal gebruikt in rapport 2. De meest waarschijnlijke verklaring van de afwijkende gemiddelden in rapport 1 zijn
  - het preferent bonusaandeel Unilever, dat zwaar doorweegt in staal 1 alhoewel het in de Legioportefeuilles weinig gewicht heeft, en dat zich erg onrepresentatief gedraagt; en
  - vermoedelijke zelfselectie in het staal: beleggers met grote posities hebben grote verliezen en zijn dus meest gemotiveerd om aan de TROS-oproep te beantwoorden; maar grote posities betekenen ook meer prijsdruk bij de uitvoering en dus slechtere koersen.<sup>11</sup> Bijgevolg heeft staal 1 gemiddeld slechtere koersen dan staal 2.

Onopgelost in rapport 1 blijft de kleine variantie aan de aankoopkant.

- Rapport 2 heeft als grote sterktes de afwezigheid van zelfselectie en de grote staalomvang. De optie, in dat rapport, om te werken met een procentuele kost aangeleverd door een externe bron is verder te verkiezen boven het werken met een (te kleine) marge uitgedrukt in cents en niet aanwijsbaar gesteund op een externe bron, wat rapport 1 doet. Onopgelost is de gunstige verkoopkoers in rapport 2.

## Bijlage: Reconstructie der standaardfouten per aandeel

De tabel vermeldt de standaarddeviaties van de gemiddelde relatieve aan- of verkoopkoersen per aandeel, uit bijlagen 4 aan rapport 2, samen met het aantal observaties. De standaardfout van de relatieve koers zelf wordt berekend als  $\sigma(\bar{X}) \times \sqrt{N}$ . Bij de verkoopkoersen zijn er twee uitschieters, waarschijnlijk niet toevallig de aandelen met weinig transacties; afgezien daarvan is er een verrassende overeenstemming. Aan de aankoopzijde is de overeenstemming even verbijsterend. Het verschil tussen de standaarddeviaties aan- v verkoop kan ik niet verklaren.

**Tabel 9: Standaardfouten per aandeel in staal 2**

	abn	aegon	ahold	fortis	ING	RDtch	tpg	unilvr	woltrs
verkoopen									
stddev( $\bar{X}$ )	1.6	1.6	2	1.6	1.6	1.9	2.4	3.1	3.6
aantal observaties	233	232	156	227	243	166	38	64	23
geïmpliceerde $\sigma_X$	24.4	24.4	25.0	24.1	24.9	24.5	14.8	24.8	17.3
aankopen									
stddev( $\bar{X}$ )	1.5		1.8	2.3	1.5	2.3			
aantal observaties	184		133	84	187	95			
geïmpliceerde $\sigma_X$	20.3		20.8	21.1	20.5	22.4			

<sup>11</sup>De aanwezigheid van prijsdruk wordt hierna statistisch aangetoond.



